Numerical Linear Algebra Days (Due giorni di Algebra Lineare Numerica)

Wednesday, 10 May 2023 - Friday, 12 May 2023

Gran Sasso Science Institute, L'Aquila

Book of Abstracts

Contents

Parallel-in-time solver for the all-at-once Runge–Kutta discretization 1

How perturbations propagate along the solutions of linear ordinary differential equations: a relative error analysis 1

A Smoothing Analysis for Multigrid Methods Applied to Tempered Fractional Problems 2

An efficient block rational Krylov solver for Sylvester equations with adaptive pole selection 2

Reduced GLT sequences and Applications 3

Structural characterization of water networks 3

Computation of the von Neumann entropy of large matrices via trace estimation and rational Krylov methods 4

Symmetrization Techniques in Image Deblurring 4

The construction of orthogonal Cauchy-like matrices 5

Fractional graph Laplacian for image reconstruction 6

A proposal for ranking through selective computation of centrality measures 6

Data-driven prediction: from LTI to NARX systems 7

Spectral analysis of weights-based finite element methods 7

Optimising the Perron eigenvalue 8

Efficient computation of the sinc matrix function for the integration of second-order differential equations 9

Dynamical low-rank training of neural networks 9

Nodal domain count for the graph p-Laplacian 10

A low rank ODE for spectral clustering stability 10

Computing closest singular matrix-valued functions 11

Kronecker-structured sketching for compressing tensors 12

A robust and conservative dynamical low-rank algorithm 12

Spectral properties of structured graphs: theory and numerical applications 13

Training of Stable Neural Ordinary Differential Equations 13

Sketched and truncated polynomial Krylov subspace methods 14

Asymptotic spectral analysis: two non-normal applications 14

1

Parallel-in-time solver for the all-at-once Runge–Kutta discretization

Author: Santolo Leveque[1]

Co-authors: Luca Bergamaschi [2]; Angeles Martinez [3]; John Pearson [4]

[1] *Scuola Normale Superiore*

[2] *University of Padova*

[3] *University of Trieste*

[4] *University of Edinburgh*

Corresponding Author: santolo.leveque@sns.it

Time-dependent PDEs arise very often in many scientific areas, such as mechanics, biology, economics, or chemistry, just to name a few. Of late, many researchers have devoted their effort to devising parallel-in-time methods for the numerical solution of time-dependent PDEs. As opposed to the classical approach, in which an approximation of the solution at a time t is computed by a sequential time-stepping, parallel-in-time methods approximate the solution of the problem for all times concurrently. This in turns adds a new dimension of parallelism and allows to speed-up the numerical solution on modern supercomputers.

In this talk, we present a fully parallelizable preconditioner for the all-at-once linear system obtained when employing a Runge–Kutta method in time. The resulting system is solved iteratively for the numerical solution and for the stages of the method. By employing classical theory of block matrices, one is able to derive an optimal preconditioner for the system considered. This results in a block-diagonal solve for all the stages at all the time-steps, and a Schur complement whose inverse can be applied by solving again for the systems for the stages of the method. Since at each linear iteration one has to solve for the latter system, we introduce a new block-preconditioner based on the SVD of the (real) Runge–Kutta coefficient matrix $A_{\mathrm{RK}} = U\Sigma V^{\top}$.

A range of numerical experiments validate the robustness of the preconditioner with respect to the discretization parameters and to the number of stages, as well as very promising scalability and parallel efficiency indices.

2

How perturbations propagate along the solutions of linear ordinary differential equations: a relative error analysis

Authors: Stefano Maset[1]; asma farooq[2]

[1] *University of Trieste, Italy*

[2] *Gran Sasso Science Institute L'Aquila Italy*

Corresponding Author: asma.farooq@gssi.it

In this talk, we are going to present how perturbations in the co-efficient matrix A propagate along the solutions of n-dimensional linear ordinary differential equations

$$\begin{cases} y'(t) = Ay(t),\ t \geq 0, \\ y(0) = y_0. \end{cases}$$

In other words we are considering the conditioning of the problem

$(\mathrm{y}_0, A) \mapsto e^{tA} y_0$

and an asymptotic analysis of condition numbers, as $t \to +\infty$, will be given. The analysis is accomplished for the case where A is normal matrix.

We remark that conditioning of such problems attained less attention in literature. At the best of our knowledge there are only two papers [1] and [2] on this topic. These papers present computational aspects of the condition number. On the other hand our study is more on theoretical aspects of the condition number. It studies how this condition number depends on the time t and the initial data y_0. Also the asymptotic behavior of condition number as $t \mapsto +\infty$ is part of our study.

[1]. A. Al-Mohy and N. Higham. Computing the action of the matrix exponential, with an application to exponential integrators. SIAM Journal on Scientific Computing, 33 (2011) no. 2, 488-511.

[2]. E. Deadman. Estimating the condition number of $f(A)b$. Numerical Algorithms 70 (2015), 287–308..

3

A Smoothing Analysis for Multigrid Methods Applied to Tempered Fractional Problems

Author: Danyal Ahmad[1]

Co-authors: Marco Donatelli [2]; Mariarosa Mazza [2]; Serra Capizzano Stefano [2]; Ken Trotti [3]

[1] *University of Insubria, Como.*

[2] *University of Insubria, Como*

[3] *University of the Italian Switzerland*

Corresponding Author: dahmad@uninsubria.it

We consider the numerical solution of time-dependent space-tempered fractional diffusion equations. The use of Crank-Nicolson in time and of second-order accurate tempered weighted and shifted Grunwald difference in space leads to dense (multilevel) Toeplitz-like linear systems. By exploiting the related structure, we design an ad-hoc multigrid solver and multigrid-based preconditioners, all with weighted Jacobi as smoother. A new smoothing analysis is provided, which refines state-of-the-art results expanding the set of suitable Jacobi weights. Furthermore, we prove that if a multigrid method is effective in the non-tempered case, then the same multigrid method is effective also in the tempered one. The numerical results confirm the theoretical analysis, showing that the resulting multigrid-based solvers are computationally effective for tempered fractional diffusion equations.

4

An efficient block rational Krylov solver for Sylvester equations with adaptive pole selection

Authors: Angelo Alberto Casulli[1]; Leonardo Robol[2]

[1] *Scuola Normale Superiore di Pisa*

[2] *Università di Pisa*

Corresponding Author: angelo.casulli@sns.it

We present an algorithm for the solution of Sylvester equations with right-hand side of low rank, based on projection onto a block rational Krylov subspace. Extending the convergence analysis in [2] to the block case, we link the convergence with the problem of minimizing the norm of a

small rational matrix over the spectra or field-of-values of the involved matrices. This is in contrast with the non-block case, where the minimum problem is scalar, instead of matrix-valued. Replacing the norm of the objective function with an easier to evaluate function yields several adaptive pole selection strategies, providing a theoretical analysis for known heuristics, as well as effective novel techniques.

[1] A. Casulli, L. Robol, An efficient block rational Krylov solver for Sylvester equations with adaptive pole selection. arXiv:2301.08103, 2023.
[2] B. Beckermann. An error analysis for rational Galerkin projection applied to the Sylvester equation. SIAM Journal on Numerical Analysis, 2011.

5

Reduced GLT sequences and Applications

Author: Giovanni Barbarino[1]

[1] *Aalto University*

Corresponding Author: giovanni.barbarino@gmail.com

The spectral analysis of matrix-sequences generated by the discretization and numerical approximation of partial differential equations, in case the domain is a generic Peano–Jordan measurable set, can be performed through the lens of generalized locally Toeplitz (GLT) theory.
In fact, it is observed that such matrix-sequences often present a spectral symbol, a measurable function describing the asymptotic behaviour of the eigenvalues.
When the domain is a hypercube, the classic GLT sequences are enough to determine the symbol, but in case of generic domains, a different kind of matrix-sequences and theory has to be formalized.
We thus develop in full detail the theory of Reduced GLT sequences and symbols, presenting some applications to finite differences and finite elements discretizations of PDEs.
In particular, we show how the theory of Reduced GLT can be exploited when the discretization grid is a mesh-up of multople grids with different degrees of fineness on a partition of the domain.

6

Structural characterization of water networks

Authors: Michele Benzi[1]; Isabella Daidone[2]; Chiara Faccio[1]; Laura Zanetti-Polzi[3]

[1] *Scuola Normale Superiore*

[2] *Department of Physical and Chemical Sciences, University of L'Aquila*

[3] *Center S3, CNR-Institute of Nanoscience, Modena*

Corresponding Author: chiara.faccio@sns.it

Liquid water, besides being fundamental for life on Earth, has long fascinated scientists due to several anomalies. Different hypotheses have been put forward to explain these peculiarities. The most accredited one foresees the presence in the supercooled region of two phases at different densities: the low-density liquid phase (LDL) and the high-density liquid phase (HDL). In a previous study [1], we proposed a new order parameter based on graph-theory, the total communicability [2], to identify these two forms in water networks at temperatures that cross the liquid-liquid coexistence line at 1950 bar. In [3], we analyzed the structure of the LDL and HDL phases at 1950 bar using consolidated global metrics from network theory. The results showed that these networks are only moderately complex, and that the lattice-like regular structure of the low-density phase is perturbed in the high-density form. More recently, we demonstrated that the total communicability can identify the two liquid phases also along the 1 bar isobar, along which there is no phase transition between LDL and

HDL, but rather a continuous change in the amount of the two components with the temperature. Furthermore, we extended the analysis conducted in [3] to these 1 bar networks, confirming the moderate complexity of these water networks. Despite the moderate complexity, the total communicability is not only able to catch the changes in the networks that characterize the two phases, but is also computationally efficient compared to other centrality measures.

[1] C. Faccio, M. Benzi, L. Zanetti-Polzi, & I. Daidone, "Low-and high-density forms of liquid water revealed by a new medium-range order descriptor", Journal of Molecular Liquids, 355, 118922, 2022.
[2] M. Benzi, & C. Klymko, "Total communicability as a centrality measure", Journal of Complex Networks, Volume 1, Issue 2, Pages 124-149, 2013.
[3] M. Benzi, I. Daidone, C. Faccio, & L. Zanetti-Polzi, "Structural analysis of water networks", Journal of Complex Networks, Volume 11, Issue 1, 2023.

7

Computation of the von Neumann entropy of large matrices via trace estimation and rational Krylov methods

Authors: Igor Simunec[1]; Michele Benzi[1]; Michele Rinelli[1]

[1] *Scuola Normale Superiore*

Corresponding Author: igor.simunec@sns.it

In this talk I will consider the approximate computation of the von Neumann entropy of a large density matrix A (i.e., a symmetric positive semidefinite matrix with unit trace), defined as $\mathrm{trace}(f(A))$, where $f(z) = -z\log(z)$, which is an important task in several fields, such as quantum information theory and network science.
The approximation of $\mathrm{trace}(f(A))$ can be reduced to the computation of a certain number of quadratic forms $x^T f(A)x$ and matrix vector products $f(A)x$, using either a probing method based on graph colorings [1] or a stochastic trace estimator [2]; in turn, these quantities can be efficiently approximated using polynomial and rational Krylov subspace methods [3].
After introducing these techniques, I will present some error bounds and heuristics for the special case of the von Neumann entropy, obtained by exploiting an integral expression of the entropy function. The analysis leads to algorithms that, given an input tolerance ϵ, aim to compute an approximation of the entropy with relative accuracy ϵ, using either theoretical bounds or heuristic estimates as stopping criteria.
The methods are tested on several density matrices from network theory to demonstrate their effectiveness.

[1] A. Frommer, C. Schimmel, and M. Schweitzer, *Analysis of probing techniques for sparse approximation and trace estimation of decaying matrix functions*, SIAM Journal on Matrix Analysis and Applications, 2021.
[2] R. A. Meyer, C. Musco, C. Musco, and D. P. Woodruff, *Hutch++: optimal stochastic trace estimation*, Symposium on Simplicity in Algorithms (SOSA), 2021.
[3] S. Güttel, *Rational Krylov approximation of matrix functions: numerical methods and optimal pole selection*, GAMM-Mitteilungen, 2013.
[4] M. Benzi, M. Rinelli, I. Simunec, *Computation of the von Neumann entropy of large matrices via trace estimation and rational Krylov methods*, arXiv preprint arXiv:2212.09642, 2022.

8

Symmetrization Techniques in Image Deblurring

Authors: Marco Donatelli[1]; Paola Ferrari[2]; Silvia Gazzola[3]

[1] *University of Insubria, Como*

[2] *Bergische Universität Wuppertal*

[3] *University of Bath*

Corresponding Author: ferrari@uni-wuppertal.de

This talk focuses on preconditioning techniques that enhance the performance of iterative regularization methods in image deblurring. The preconditioners are applied to problems with different point spread functions (PSFs) and boundary conditions [1]. More precisely, we first consider the anti-identity preconditioner [3], which symmetrizes the coefficient matrix associated to problems with zero boundary conditions, allowing the use of MINRES as a regularization method. When considering more sophisticated boundary conditions and strongly nonsymmetric PSFs, the anti-identity preconditioner improves the performance of GMRES. We present both stationary and iteration-dependent regularizing circulant preconditioners that speed up the iterations when applied in connection with the anti-identity matrix and flexible Krylov subspaces [2]. A theoretical result about the clustering of the eigenvalues of the preconditioned matrices is proved in a special case [1]. The results of many numerical experiments are illustrated to show the effectiveness of the new preconditiong techniques, including when considering the deblurring of sparse images.

[1] M. Donatelli, P. Ferrari, S. Gazzola, *Symmetrization Techniques in Image Deblurring*, https://arxiv.org/abs/2212.05879

[2] S. Gazzola, J. G. Nagy, M. Sabaté Landman, *Iteratively reweighted FGMRES and FLSQR for sparse reconstruction*, SIAM J. Sci. Comput., 2021.

[3] A. J. Wathen, J. Pestana, *A preconditioned MINRES method for nonsymmetric Toeplitz matrices*, SIAM J. Matrix Anal. Appl., 2015.

9

The construction of orthogonal Cauchy-like matrices

Author: Dario Fasino[1]

[1] *University of Udine*

Corresponding Author: dario.fasino@uniud.it

A matrix $K \in \mathbb{R}^{n\times n}$ is Cauchy-like if its entries have the form

$$\mathrm{K}_{ij} = \frac{a_i b_j}{x_i - y_j}, \qquad i,j = 1,\dots,n,$$

where x_i, y_j for $i,j = 1,\dots,n$ are pairwise distinct real numbers. Besides their pervasive occurrence in computations with rational functions, Cauchy-like matrices play an important role in deriving algebraic and computational properties of many displacement-structured matrix families and occur as fundamental blocks (together with trigonometric transforms) in decomposition formulas and fast solvers for, e.g., Toeplitz, Hankel, and related matrices.

This contribution provides a complete description of orthogonal Cauchy-like matrices [1]. Interest in these matrices stems from the paper [2], where they are needed to design all-pass filters for signal processing, and a novel characterization of Cauchy matrices as transition matrices between eigenbases of particular matrix pairs [3]. We illustrate their relationships with secular equations, the diagonalization of symmetric quasi-separable matrices and the construction of orthogonal rational functions with free poles. Moreover, we characterize matrix algebras that are simultaneously diagonalized by orthogonal Cauchy-like matrices.

[1] D. Fasino, Orthogonal Cauchy-like matrices, Numerical Algorithms, 92 (2023), 619–637.

[2] S. J. Schlecht, Allpass feedback delay networks, IEEE Trans. Signal Process., 69 (2021), 1028–1038.

[3] A. G. Lynch, Cauchy pairs and Cauchy matrices, Linear Algebra Appl., 471 (2015), 320–345.

10

Fractional graph Laplacian for image reconstruction

Author: Marco Donatelli[1]

Co-authors: Alessandro Buccini [2]; Stefano Aleotti [3]

[1] *University of Insubria, Como*

[2] *Università di Cagliari*

[3] *Università degli studi dell'Insubria*

Corresponding Author: saleotti@uninsubria.it

Image reconstruction problems, like image deblurring and computer tomography, are usually ill-posed and require regularization.
A popular approach to regularization is to substitute the original problem with an optimization problem that minimizes the sum of two terms, an ℓ^2 term and an ℓ^q term with $0 < q \leq 1$. The first penalizes the distance between the measured data and the reconstructed one, the latter imposes sparsity on some features of the computed solution.

In this work, we propose to use the fractional Laplacian of a properly constructed graph in the ℓ^q term to compute extremely accurate reconstructions of the desired images.
A simple model with a fully plug-and-play method is used to construct the graph and enhanced diffusion on the graph is achieved with the use of a fractional exponent in the Laplacian operator. Since this is a global operator, we propose to replace it with an approximation in an appropriate Krylov subspace.

[1] D. Bianchi, A. Buccini, M. Donatelli, E. Randazzo,
"Graph Laplacian for image deblurring", Electronic Transactions on Numerical Analysis, 2021, 55, pp. 169-186.
[2] A. Buccini, M. Donatelli, "Graph Laplacian in $\ell^2 - \ell^q$ regularization for image reconstruction", Proceedings - 2021 21st International Conference on Computational Science and Its Applications, ICCSA 2021, 2021, pp. 29-38.
[3] S. Aleotti, A. Buccini, M. Donatelli, "Fraction Graph Laplacian for image reconstruction", in progress.

11

A proposal for ranking through selective computation of centrality measures

Authors: Alessandro Filippo[1]; Daniele Bertaccini[1]

[1] *Università degli studi di Roma Tor Vergata*

Corresponding Author: filippo@mat.uniroma2.it

In complex network analysis it is crucial to investigate the alteration of network structures that results from the targeted removal of vertices or edges, ranked by a centrality measure [3]. Unfortunately, a sequential recalculation of centralities after each node elimination is often impractical for large networks, and using the ranking before the removal often does not approximate accurately the actual scenario.

In [2] we propose a first result on the computational complexity of the sequential approach when considering percolation on a complex network using some centrality measures based on matrix functions [1]. Moreover, we present two strategies that aim to reduce the computational impact of the sequential computation of centralities and provide theoretical results in support. Finally, we provide an application of our claims to the robustness of some synthetic and real-world networks using the same approach as [4].

We will also present an attempt to approximate updating of some centrality measures based on matrix functions to extend our proposal in [2].

[1] M. Benzi and P. Boito, Matrix Functions in Network Analysis. GAMM-Mitteilungen, 2020.
[2] D. Bertaccini and A. Filippo, A proposal for ranking through selective computation of centrality measures, 2023, in revisione.
[3] E. Estrada, The Structure of Complex Networks: Theory and Applications, Oxford University Press, 2011.
[4] S. Iyer, T. Killingback, B. Sundaram and Z. Wang, Attack Robustness and Centrality of Complex Networks, PLOS ONE, 2013.

12

Data-driven prediction: from LTI to NARX systems

Authors: Alessandro Chiuso[1]; Antonio Fazzi[1]

[1] *University of Padova*

Corresponding Author: antonio.fazzi@unipd.it

The behavioral setting [1] is suited for data-driven algorithms since systems are viewed as sets of trajectories. A classical result in this framework, known as /Willems' fundamental lemma/ [2], states the conditions that allow to represent all the system trajectories from an observed one. This result makes possible to perform data-driven simulations [3], that is simulation of the future system trajectories directly from the observed data (without estimating a system model).
The classical theory about the topic was developed for the class of linear time-invariant systems only.
We discuss recent results [4,5] on how to switch from linear to nonlinear systems.

[1] J. W. Polderman and J. C. Willems. Introduction to Mathematical Systems Theory, volume 26 of Texts in Applied Mathematics. Springer New York, New York, NY, 1998.

[2] J. C. Willems, P. Rapisarda, I. Markovsky, and B. De Moor. A note on persistency of excitation. Syst. Control Lett., 54(4):325–329, 2005.

[3] I. Markovsky and P. Rapisarda, "Data-driven simulation and control," Int. J. Control, vol. 81, pp. 1946–1959, 2008.

[4] I. Markovsky. Data-driven simulation of generalized bilinear systems via linear time-invariant embedding. IEEE Trans. Automat. Contr., 2023.

[5] A. Fazzi and A. Chiuso. Data-driven prediction and control for NARX systems. Submitted.

13

Spectral analysis of weights-based finite element methods

Authors: Ludovico Bruni Bruno[1]; Matteo Semplice[1]; Stefano Serra Capizzano[1]

[1] *Università dell'Insubria*

Corresponding Author: ludovico.brunibruno@uninsubria.it

A finite element method is defined by specifying three objects: a mesh, a finite dimensional functional space of shape functions and a collection of degrees of freedom, namely linear functionals on shape functions. It is rather this third aspect that characterises the method.

In the last twenty years a large amount of research has been devoted to generalising existing methods. These extensions moved towards several directions: flexible frameworks, more abstract formulations involving differential forms and, of course, several choices of alternative degrees of freedom have been consequently proposed. Among them we recall, for instance, *moments* and *weights*. A weight is, in its greatest generality, the integral of a multilinear k-form (tipically a differential form) on a k-simplex. As a consequence, this latter choice naturally generalises Lagrangian finite elements, allowing for considering physical measurements as degrees of freedom.

In practice, changing degrees of freedom consists in choosing a different basis for the bilinear form of the problem on the discrete level, from any given basis to that in (a fixed) duality with the specified degrees of freedom. Any different basis thus yields a different stiffness matrix and hence a different spectral symbol, which is, in the case of finite elements, a $r \times r$ Hermitian matrix-valued function, being $r + 1 = R$ the dimension of the local space of shape functions. For a d-variate problem, the resulting matrices have essentially the structure of d-level $R \times R$ positive definite Toepliz matrices.

Although the dimension of the stiffness matrix of the problem grows linearly with the number of elements, the rank of the symbol does not change and such a function perfectly captures the spectral behaviour of the whole stiffness matrix. As a consequence, this tool has been used to analyse and optimise several approximation methods, such as Lagrangian finite elements. We rely on this function for a spectral analysis of finite elements based on non-standard degrees of freedom. In particular, we focus on the Laplacian operator and offer a spectral analysis of weights, seeking for optimal distributions of supports in terms of the conditioning of the corresponding stiffness matrix. Results are compared with the well established counterpart for classical Lagrangian elements.

[1] A. Alonso Rodríguez, L. Bruni Bruno and F. Rapetti, Towards nonuniform distribu- tions of unisolvent weights for high-order Whitney edge elements, Calcolo, 59 (2022).
[2] S. H. Christiansen and F. Rapetti, On high order finite element spaces of differential forms, Math. Comp., 85 (2016), 517–548.
[3] D. N. Arnold, R. S. Falk, R. Winther, Finite element exterior calculus, homological techniques, and applications, Acta Numerica, 15 (2006), 1–155.
[4] C. Garoni, S. Serra Capizzano and D. Sesana, Spectral analysis and spectral symbol of d-variate $\mathbb{Q}_p$ Lagrangian FEM stiffness matrices, SIAM J. Mat. Anal. App., 36 (2015), 1100– 1128.

14

Optimising the Perron eigenvalue

Author: Vladimir Protasov[1]

[1] *University of L'Aquila*

Corresponding Author: vladimir.protasov@univaq.it

The problem of optimizing of the spectral radius over a convex set of matrices is notoriously hard. The spectral radius is neither convex nor concave. It can be non-Lipschitz either, which makes the standard gradient-type methods inapplicable. We consider the case of nonnegative matrices, when the spectral radius becomes the Perron eigenvalue. For the matrix sets with the product structure (with row or column uncertainty), we consider the spectral simplex method and the greedy method [1,2], which demonstrate a surprising efficiency even in high dimensions (several thousands). For polyhedra of nonnegative matrices, we present and discuss an alternating algorithm. Applications to the graph theory, the neural search, and the population dynamics are discussed.

[1] Protasov, V.Y., Spectral simplex method
Mathematical Programming, 2016, 156 (1-2), pp. 485–511

[2] Cvetković, A., Protasov, V.Yu. The greedy strategy for optimizing the Perron eigenvalue, Mathematical Programming, 2022, 193(1)

15

Efficient computation of the sinc matrix function for the integration of second-order differential equations

Authors: Fabio Durastante[1]; Lidia Aceto[2]

[1] *Department of Mathematics, University of Pisa*

[2] *Dipartimento di Scienze e Innovazione Tecnologica, Università del Piemonte Orientale*

Corresponding Author: fabio.durastante@unipi.it

In this poster [1], we deal with the numerical solution of systems of oscillatory second-order differential equations which often arise from the space semi-discretization of partial differential equations. Since these differential equations exhibit (pronounced or highly) oscillatory behavior, standard numerical methods are known to perform poorly. Our approach consists in directly discretizing the problem by employing Gautschi-type integrators [2] based on sinc matrix functions. The novelty contained here is that of using a suitable rational approximation formula for the sinc matrix function to apply a rational Krylov-like approximation method with suitable choices of poles. In particular, we discuss the application of the whole strategy to a finite element discretization of the wave equation.

[1] L. Aceto, F. Durastante. Efficient computation of the sinc matrix function for the integration of second-order differential equations, In preparation,
[2] W. Gautschi, Numerical integration of ordinary differential equations based on trigonometric polynomials, Numer. Math. 3 (1961) 381–397.

16

Dynamical low-rank training of neural networks

Authors: Emanuele Zangrando[1]; Francesco Tudisco[2]

[1] *Gran Sasso Science Institute*

[2] *GSSI*

Corresponding Author: emanuele.zangrando@gssi.it

Neural networks have achieved tremendous success in a wide variety of applications. However, their memory footprint and computational demand can render them impractical in application settings with limited hardware or energy resources. At the same time, overparameterization seems to be necessary in order to overcome the highly non-convex nature of the training optimization problem. An optimal trade-off is then to be found in order to reduce networks' dimension while mantaining high performance.
Popular approaches in the current literature are based on pruning techniques that look for subnets capable of maintaining approximately the initial performance. Nevertheless, these techniques often are not able to reduce the memory footprint of the training phase.
In this talk we will present DLRT, a training algorithm that looks for "low-rank lottery tickets" by interpreting the training phase as a continuous ODE and by integrating it within the manifold of low-rank matrices.

These subnetworks and their ranks are determined and adapted already during the training phase, allowing the overall time and memory resources required by both training and evaluation phases to be reduced significantly.
The talk is based on [1].

[1] S. Schotthöfer, E. Zangrando, J. Kusch, G. Ceruti, F. Tudisco,
"Low-rank lottery tickets: finding efficient low-rank neural networks via matrix differential equations", NeurIPS, 2022.

17

Nodal domain count for the graph p-Laplacian

Author: Piero Deidda[1]

Co-authors: Francesco Tudisco [2]; Mario Putti [3]

[1] *SNS - GSSI*

[2] *GSSI*

[3] *University of Padova*

Corresponding Author: piero.deidda@sns.it

The nodal domain count of the graph p-Laplacian eigenfunctions allows to relate the p-Laplacian eigenvalues to different topological invariants of the graph, i.e. the Cheeger constants and the packing radii of the graph. We prove that the nodal domain count of any eigenfunction can be bounded, both from above and below, in terms of the position of the corresponding eigevalue in the variational spectrum. To this end, we prove that the variational spectrum of the p-Laplacian operator on forests exhaust the spectrum, we transfer the Weyl's inequalities for the Laplacian matrix to the nonlinear p-Laplacian and we provide a Perron-Frobenious-like characterization of the first eigenpair of the p-Laplacian. Our new results show that the variational p-Laplacian eigenvalues on trees equal the Cheeger constants and packing radii of the graph, respectively for $p = 1$ and $p = \infty$. Moreover, when applied to the linear case $p = 2$, the new results imply well-known properties of the linear Laplacian matrix as well as novel ones.

References:
[1] T. Bhuler, M. Hein, Spectral clustering based on the graph p-Laplacian, Proceedings of the 26th Int. Conf. on Machine Learning (2009).
[2] K.C. Chang, Spectrum of the 1-Laplacian and Cheeger's constant on graphs, Journal of Graph Theory (2016).
[3] K.C. Chang, S. Shao, and D. Zhang, Nodal domains of eigenvectors for
1-Laplacian on graphs, Advances in Mathematics (2017).
[4] F. Tudisco, M. Hein, A nodal domain theorem and a higher-order Cheeger inequality for the graph p-Laplacian, Journal of Spectral Theory (2018).
[5] P. Deidda, The graph p-Laplacian eigenvalue problem, PhD Thesis (2023).
[6] P. Deidda, M. Putti and F. Tudisco, Nodal domain count for the generalized graph p-Laplacian, Applied and Computational Harmonic Analysis (2023).

18

A low rank ODE for spectral clustering stability

Authors: Nicola Guglielmi[1]; Stefano Sicilia[1]

[1] *Gran Sasso Science Institute*

Corresponding Author: stefano.sicilia@gssi.it

Spectral clustering is a well-known technique which identifies k clusters in an undirected graph with weight matrix $W \in \mathbb{R}^{n\times n}$ by exploiting its graph Laplacian
L(W)=diag(W**1**)-W, **1**=(1,...,1)$^T \in \mathbb{R}^n$,$whoseeigenvalues$0=$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ and eigenvectors are related to the k clusters. Since the computation of λ_{k+1} and λ_k affects the reliability of this method, the k-th spectral gap $\lambda_{k+1} - \lambda_k$ is often considered as a stability indicator. This difference can be seen as an unstructured distance between $L(W)$ and an arbitrary symmetric matrix $L_\star$ with vanishing k-th spectral gap.

A more appropriate structured distance to ambiguity such that $L_\star$ represents the Laplacian of a graph has been proposed in [1]. Slightly differently, we consider the objective functional $\mathrm{F}(\Delta) = \lambda_{k+1}\left(L(W+\Delta)\right) - \lambda_k\left(L(W+\Delta)\right)$,$where\Delta$ is a perturbation such that $W+\Delta$ has non-negative entries and the same pattern of W. We look for an admissible perturbation $\Delta_\star$ of smallest Frobenius norm such that $F(\Delta_\star) = 0$.

In order to solve this optimization problem, we exploit its low rank underlying structure. Similarly to [2], we formulate a rank-4 symmetric matrix ODE whose stationary points are the optimizers sought. The integration of this equation benefits from the low rank structure with a moderate computational effort and memory requirement, as it is shown in some illustrative numerical examples.

[1] E. Andreotti, D. Edelmann, N. Guglielmi, C. Lubich, Measuring the stability of spectral clustering, Linear Algebra and its Applications, 2021

[2] N. Guglielmi, C. Lubich, S. Sicilia, Rank-1 matrix differential equations for structured eigenvalue optimization, arXiv preprint arXiv:2206.09338, 2022

19

Computing closest singular matrix-valued functions

Authors: Miryam Gnazzo[1]; Nicola Guglielmi[1]

[1] *Gran Sasso Science Institute*

Corresponding Author: miryam.gnazzo@gssi.it

Given a set of matrices $A_i \in \mathbb{C}^{n\times n}$ and a set of analytic functions $f_i : \mathbb{C} \mapsto \mathbb{C}$, we consider a regular matrix-valued function $\mathcal{D}(\lambda) = \sum_{i=0}^{d} f_i(\lambda) A_i$, that is such that $\det\left(\mathcal{D}(\lambda)\right)$ is not identically zero for $\lambda \in \mathbb{C}$. This class of functions includes for example characteristic functions associated to linear systems of delay differential equations with constant delays. An interesting problem consists in the computation of the nearest singular function $\widetilde{\mathcal{D}}(\lambda) = \sum_{i=0}^{d} f_i(\lambda)\left(A_i + \Delta A_i\right)$, where the distance is measured in the Frobenius norm. For example, for a system of delay differential equations

$$\mathrm{A}_0\dot{x}(t) = A_1 x(t) + \sum_{i=2}^{d} A_d x(t-\tau_d)$$

the characteristic equation is $\det\left(-\lambda A_0 + A_1 + e^{-\tau_2\lambda}A_2 + \ldots + e^{-\tau_d\lambda}A_d\right) = 0$. Differently from the case of matrix polynomials, in the general case of entire functions $f_i(\lambda)$ like the one above, we have to take into account that the function $\det\left(\mathcal{D}(\lambda)\right)$ may have an infinite number of roots. This represents a delicate feature of the problem and requires an appropriate analysis in the construction of the numerical method for the computation of the distance to singularity. The condition of singularity is associated to the property that the function $\det\left(\widetilde{\mathcal{D}}(\lambda)\right)$ vanishes on a suitable closed complex curve Γ.

We propose a two level procedure, following the idea introduced in [1], and impose that the determinant vanishes on a finite set of prescribed complex points $\{\mu_j\}_{j=1}^{m}$, suitably sampled on the curve Γ. This can be translated into the minimization of the functional $\mathrm{F}_\varepsilon\left(\Delta A_0, \ldots, \Delta A_d\right) = \frac{1}{2}\sum_{j=1}^{m} \sigma_{\min}^2\left(\widetilde{\mathcal{D}}(\mu_j)\right)$,$where\sigma_{\min}$ denotes the smallest singular value and $\varepsilon = \left\|[\Delta A_0, \ldots, \Delta A_d]\right\|_F$

is the norm of the perturbation, and in finding the smallest value $\varepsilon^\star$ such that the functional vanishes.
This approach can be extended to situations where the matrix-valued functions present a certain structure. For instance, we can include in the method additional constraints, such as a certain sparsity pattern determined by the original matrices or even structures involving the whole functions, like palindromic properties.

[1] M. Gnazzo, N. Guglielmi, Computing the closest singular matrix polynomial, arXiv preprint arXiv:2301.06335, 2023.

20

Kronecker-structured sketching for compressing tensors

Authors: Alberto Bucci[1]; Leonardo Robol[2]

[1] *University of Pisa*

[2] *Università di Pisa*

Corresponding Author: alberto.bucci@phd.unipi.it

We consider the problem of performing low-rank compression through structured sketching for tensors.
In particular, we analyze the theoretical properties of sketchings formed as Kronecker or Khatri-Rao product of random embeddings.
We present new versions of the matrix Chernoff inequality, designed to provide bounds on tensorial structured matrices.
When applied to vectors with Kronecker structure these matrices dramatically reduce the embedding complexity, however in worst case scenario they may lead to very poor embeddings; as we show in the experiments.
For this reason we look for embeddings which preserve the advantages of the tensorial structured ones and which are also Johnson Lindenstrauss transformations.
We conclude by showing that a similar embedding can be obtained by slightly modifying the Kronecker and Khatri-Rao embeddings descripted above.

21

A robust and conservative dynamical low-rank algorithm

Authors: Alexander Ostermann[1]; Carmen Scalone[2]; Lukas Einkemmer[1]

[1] *University of Innsbruck*

[2] *University of L'Aquila*

Corresponding Author: carmen.scalone1990@gmail.com

In this talk, we propose a numerical strategy to solve the Vlasov-Poisson equation, based on the dynamical low rank approximation, which is extremely efficient to reduce the complexity of such model.
Previous approaches based on dynamical low rank approximation destroyed the physical structure of the problem, as physical invariants were lost.
Starting from the conservative continuous setting proposed in [1], we introduce a suited modification of the algorithm [2], able to conserve mass and momentum (up to machine precision) and significantly improves energy conservation. The main ingredients are the adding suited basis functions to the approximation space and conservative rank truncation. Finally, rank adaptation techniques are introduced. This is a joint work with L. Einkemmer and A. Ostermann.

[1] L. Einkemmer and I. Joseph, A mass, momentum, and energy conservative dynamical low-rank scheme for the Vlasov equation, J. Comput. Phys., 2021.
[2] G. Ceruti and C. Lubich, An unconventional robust integrator for dynamical low-rank approximation, BIT Numer. Math., 2022.

22

Spectral properties of structured graphs: theory and numerical applications

Author: Andrea Adriani[1]

[1] *Università dell'Insubria*

Corresponding Author: aadriani@uninsubria.it

The study of spectral properties of structured graphs constitutes a recent and fast growing research field, useful in various branches of applied mathematics.
Spectral properties of the graph-Laplacian turn out to be useful in the discretization and in the computation of solutions of partial differential equations (PDEs), in the study of spectral gaps, clustering, etc. ([1],[2],[3]).
In particular, we will focus on the study of sequences of graphs with a uniform local structure (Toeplitz graphs and their generalizations) and the spectral properties of such graphs and of the associated graph-Laplacians.
Such graphs often appear when modeling real-world problems, particularly when dealing with the discretization of PDEs. The knowledge of the spectral properties of these graphs and of their associated graph-Laplacian can give valuable insights for various numerical applications, ranging from the study of spectral gaps and clustering to the construction of optimal preconditioners ([1],[2],[3],[4]).
Our main theoretical result is giving the spectral distribution for the sequence of graph-Laplacians associated with a sequence of d-level diamond Toeplitz graphs ([1],[3]). After this, we will show some numerical applications related to the solutions of linear systems involving graph-Laplacians of structured graphs which confirm the usefulness of the developed theory ([1],[2]).

[1] A. Adriani, "Geometric, Stochastic and Spectral properties of Graphs: theory and numerical applications", PhD Thesis, 2022.

[2] A. Adriani, D. Bianchi, P. Ferrari, S. Serra-Capizzano, "Asymptotic spectra of large (grid) graphs with a uniform local structure (part II): numerical applications", Submitted, 2021.

[3] A. Adriani, D. Bianchi, S. Serra-Capizzano, "Asymptotic spectra of large (grid) graphs with a uniform local structure (part I): theory", Milan J. Math. 88: 409–454, 2020.

[4] C. Garoni, S. Serra-Capizzano, "The theory of Generalized Locally Toeplitz sequences: theory and applications - Vol I", Springer - Springer Monographs in Mathematics, New York, 2017.

23

Training of Stable Neural Ordinary Differential Equations

Authors: Arturo De Marinis[1]; Nicola Guglielmi[1]; Anton Savostianov[1]; Francesco Tudisco[2]; Emanuele Zangrando[1]

[1] *Gran Sasso Science Institute*

[2] *GSSI*

Corresponding Author: arturo.demarinis@gssi.it

Neural ordinary differential equations (neural ODEs) are a new family of deep neural networks. Essentially, a neural ODE is a differential equation whose vector field is a neural network.
Having a neural ODE as a part of a machine learning model makes the model more efficient than a standard one. Indeed, it is possible to train the neural ODE block of the model using the adjoint sensitivity method, which computes the gradients for the gradient descent method avoiding the computational cost of the classical backpropagation.
Our contribution to this field is the study of the stability and the contractivity of the neural ODE block, being a differential equation, with the aim of designing training strategies to make the overall machine learning model robust and stable against adversarial attacks.
This poster is based on [1], [2] and [3].

[1] Ricky T. Q. Chen, Yulia Rubanova, Jesse Bettencourt and David K. Duvenaud, Neural Ordinary Differential Equations, Advances in Neural Information Processing Systems 31, 2018.
[2] Eldad Haber and Lars Ruthotto, Stable Architectures for Deep Neural Networks, Inverse Problems 34(1), 2018.
[3] Ian J. Goodfellow, Jonathon Shlens and Christian Szegedy, Explaining and Harnessing Adversarial Examples, International Conference on Learning Representations, 2015.

24

Sketched and truncated polynomial Krylov subspace methods

Author: Davide Palitta[1]

Co-authors: Marcel Schweitzer [2]; Valeria Simoncini [1]

[1] *Alma Mater Studiorum, Università di Bologna*

[2] *Bergische Universität Wuppertal*

Corresponding Author: davide.palitta@unibo.it

Sketching can be seen as a random dimensionality reduction able to preserving the main features of the original problem with probabilistic confidence. Such kind of techniques is emerging as one of the most promising tools to boost numerical computations and it is quite well-known by theoretical computer scientists. Nowadays, sketching is getting popularity also in the numerical linear algebra community even though its use and understanding are still limited.
In this talk we present the main concepts related to sketching and how the latter can be combined with Krylov subspace methods.
We will focus on the solution of large-scale linear systems as model problem. On the other hand, thanks to the novel sketched Arnoldi relation we will illustrate, the results discussed in this talk can be extended to a panel of diverse algebraic problems ranging from the numerical evaluation of matrix functions to the solution of matrix equations.

This talk is based on a joint work with Valeria Simoncini and Marcel Schweitzer.

25

Asymptotic spectral analysis: two non-normal applications

Authors: Alec Jacopo Almo Schiavoni Piazza[1]; David Meadon[2]; Stefano Serra-Capizzano[1]

[1] *University of Insubria*

[2] *Uppsala University*

Corresponding Author: ajapiazza@studenti.uninsubria.it

The study of the asymptotic properties of the spectrum of a matrix sequence with a specific structure and increasing order has been a rich field of research and of great interest in applications, since they arise from approximation of integro-differential equations, also of fractional order (see e.g. [4, 5, 3] and references therein). This kind of approach leads most of the times to a huge linear system, whose dimension is related to the precision of the approximation. For this reason, researchers focused their attention on properties of clustering and symbol analysis of the eigenvalues of the sequence, since those are connected with fast convergence of iterative procedures for solving linear systems, such as Krylov-type methods [2, 6]. While classical results are often efficient in the case of a sequence of normal matrices, a wilder behavior should be expected in non-normal settings. In this poster, two such cases [9, 10] are presented and specific tools as long as complex analysis and quite new literature are involved to obtain the main results.

[1] G. Barbarino, S. Serra-Capizzano. Non-Hermitian perturbations of Hermitian matrix-sequences and applications to the spectral analysis of the numerical approximation of partial differential equations. Numer. Linear Algebra Appl. 27 (2020), no. 3, e2286, 31 pp.
[2] B. Beckermann, A.B.J. Kuijlaars. Superlinear convergence of conjugate gradients. SIAM J. Numer. Anal. 39-1 (2001), 300–329.
[3] M. Donatelli, M. Mazza, S. Serra-Capizzano. Spectral analysis and preconditioning for variable coefficient fractional derivative operators. J. Comput. Phys. 307 (2016), 262–279.
[4] C. Garoni, S. Serra-Capizzano. The theory of Generalized Locally Toeplitz sequences: theory and applications - Vol I. SPRINGER - Springer Monographs in Mathematics, Berlin, (2017).
[5] C. Garoni, S. Serra-Capizzano. The theory of Generalized Locally Toeplitz sequences: theory and applications - Vol II. SPRINGER - Springer Monographs in Mathematics, Berlin, (2018).
[6] A.B.J. Kuijlaars. Convergence analysis of Krylov subspace iterations with methods from potential theory. SIAM Rev. 48-1 (2006), 3–40.
[7] D.S. Lubinsky. Distribution of eigenvalues of Toeplitz matrices with smooth entries. Linear Algebra Appl. 633 (2022), 332–365.
[8] W. Parry. On the β-expansions of real numbers. Acta Math. Acad. Sci. Hungar. 11 (1960), 401–416.
[9] A.J.A. Schiavoni-Piazza, S. Serra-Capizzano. Distribution results for a special class of matrix sequences: joining approximation theory and asymptotic linear algebra. Electron. Trans. Numer. Anal. 59 (2023), 1–8
[10] A.J.A. Schiavoni-Piazza, D. Meadon, S. Serra-Capizzano. The β maps and the strong clustering at the complex unit circle. Submitted for publication.
[11] J.-L. Verger-Gaugry. Uniform distribution of Galois conjugates and beta-conjugates of a Parry number near the unit circle and dichotomy of Perron numbers. Unif. Distrib. Theory 3-2 (2008), 157–190.